**Serie storiche**

**Materiale trovato online:**

When investigating a time series, one of the first things to check before building an ARIMA model is to check that the **series is stationary**. That is, it needs to be determined that the time series is constant in mean and variance are constant and not dependent on time.

→ Nel nostro caso, i due test non concordano

Before running any models, plot your time series data and take a good look at it. If there is a trend and/or seasonality, go for exponential smoothing methods (ETS) because the ETS family explicitly model these components. In case you see autocorrelation in the data, i.e. the past explains the present, go for the ARIMA methodology. Plotting the partial autocorrelation function can help you with this.

Mentally choose a preferred method based on the method described above but run both ARIMA and ETS models. You can choose the best method based on validation methods like out-of-sample validation or information criteria like AIC and BIC. These methods generally try to estimate the real SSE of your method and help you choose the ‘best’ method.

**Procedura generale:**

Exponential smoothing methods are useful for making forecasts, and make no assumptions about the correlations between successive values of the time series. However, if you want to make prediction intervals for forecasts made using exponential smoothing methods, the prediction intervals require that the forecast errors are uncorrelated and are normally distributed with mean zero and constant variance.

**ARIMA: tiene conto delle autocorrelazioni**

**ETS: tiene conto di trend, errore e componente stagionale**

**N.B.: i dati vanno convertiti in oggetto time series univariata per poter applicare i modelli**

In ogni caso, dopo aver fatto un modello di previsione, occorre fare check sul fatto che gli errori di previsione non siano correlati, abbiano varianza costante e media nulla.

NOSTRO CASO:

Fattore diceva che se considerassimo solo i regressori esterni, non terremmo in considerazione l’autoregressione e i dati dei mesi precedenti ma considereremmo solo una previsione basata su 0-1.

Secondo me invece si potrebbe provare a valutare la nostra idea iniziale (dovremmo spostare sia ore di luce e ore lavorative) e vedere come va, usando regressori esterni. Se poi capiamo come implementare anche la parte di fattore ci possiamo provare ma non esageriamo.

Come fare regressori esterni:

* auto.arima permette regressori esterni con xreg, ripetendo poi anche in forecast
* ets non permette di farlo, nemmeno tbats; posso sfruttare auto.arima con xreg con covariate e serie di fourier (tiene conto della stagionalità multipla), lo svantaggio è che arima forza la stagionalità ad essere periodica e dunque non dinamica come permetterebbe tbats da solo.

<https://otexts.com/fpp2/complexseasonality.html>

<https://robjhyndman.com/hyndsight/tbats-with-regressors/>

Articoli interessanti per idea di fattore:

* <https://arxiv.org/pdf/2002.05746.pdf> (considera un cambiamento di policy governativa in una città e vuole prevedere quale sarebbe stato il comportamento/la situazione se la policy non fosse cambiata, considera diversi approcci/modelli, nel nostro caso potrebbe essere policy0=doppio orario, policy1=singolo orario, vogliamo prevedere cosa succederebbe cambiando policy (task opposto, non so se funzionerebbe comunque)).
* <https://arxiv.org/pdf/1706.07840.pdf> stima causal effects e da lì ricava potential outcome alternativo a quello osservato. Propone diversi metodi per fare la stima a seconda del problema in studio, bisognerebbe leggerli tutti bene.

Giornate con valori nulli: potrei “creare” dati con analisi valori vicini ma non li userei come test, solo nel training. Posso provare a trasformare gli 0 in na e vedere se il modello riesce a tenerli in considerazione correttamente.